This is the peer reviewed version of the following article: Allen C, Havlíček J, Williams K, Roberts SC. Perfume experts' perceptions of body odors: Toward a new lexicon for body odor description. *Journal of Sensory Studies* 2018;33:e12314, which has been published in final form at https://doi.org/10.1111/joss.12314. This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving.

Accepted Final Version – Journal of Sensory Studies.

First published online 22 January 2018

Perfume experts' perceptions of body odours: Towards a new lexicon for body odour description

Running title: A new lexicon for body odour description

Allen, C.[1,2*], Havlíček, J.[3], Williams, K.[4†], & Roberts, S.C.[2]

1 School of Psychology, Newcastle University, Newcastle-upon-Tyne, NE1 7RU, UK

2 Division of Psychology, University of Stirling, FK9 4LA, UK

3 Department of Zoology, Charles University, Prague, Czech Republic

4 Seven Scent Ltd, Lamplight Way, Manchester, M27 8UJ, UK

* Correspondence to be sent to: Dr Caroline Allen, Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK. Email: caroline.allen@newcastle.ac.uk

†Current address: Unilever Research and Development, Port Sunlight Laboratory, Wirral, UK

Funding

JH is supported by the Czech Science Foundation grant [GA15-05048S].

Abstract Human axillary (armpit) odours are highly diverse and have potential to reveal a wide range of individual information. This is echoed in gas chromatography findings, which show that axillary odours are comprised of many volatile compounds. Despite this, only a small number of verbal descriptors are used when investigating the perceptual qualities of body odours. We set out to develop a lexicon that would capture these perceptual qualities in more detail, working alongside perfumers and fragrance evaluators in order to benefit from their expertise in olfactory perception and semantic labelling of odours. Four experts developed a list of 15 verbal descriptors based on an exemplar set of male and female axillary samples, and then rated 62 samples (31 men and 31 women) using these. We explored the predictive value of these ratings, finding that subsets of descriptors distinguished male from female samples, appearing to be more reliable than explicit judgments of odour sex.

Practical applications This lexicon was successful in discriminating sex of odour samples and could enable improved understanding of other perceptual qualities of human odour. For example, it could be possible to link specific perceptual qualities to specific cues (e.g. symmetry, masculinity) or to manipulate odours based on perceptual qualities in experimental settings, with direct practical implications for odour researchers. Furthermore, the existence of such a lexicon will allow body odours to be categorised for practical purposes. For example, such categorisation will facilitate exploration of how fragrances, ingredients or accords may interact with and complement different body odour types.

Keywords: Odor classification, Olfaction, Olfactory perception, Sex identification, Smell, Verbal descriptors

Introduction

Human odours are multi-faceted, as reflected by the range of information which appears to be detectable by conspecifics, from stable traits such as genetic information (Havlíček & Roberts, 2009; Roberts et al., 2005; Wedekind, Seebeck, Bettens, & Paepke, 1995; Winternitz, Abbate, Huchard, Havlíček, & Garamszegi, 2017) and developmental stability (Rikowski & Grammer, 1999) through to those which fluctuate such as emotions (Chen & Haviland-Jones, 2000; Sorokowska, Sorokowski, & Szmajke, 2012) health (Moshkin et al., 2012), diet (Fialová, Roberts, & Havlíček, 2016; Havlicek & Lenochova, 2006) and fertility status (Havlíček, Dvořáková, Bartoš, & Flegr, 2006; Kuukasjärvi et al., 2004). In line with this diversity, human axillary odours are comprised of hundreds of volatile compounds, some of which appear to be sex- or individual-specific, potentially indicating genetic information (Penn et al., 2007).

Despite the variety of socially relevant cues which appear to be present and assessable in odours, most studies to date employ simple and, arguably, vague terminology when asking participants to rate odour samples. Most commonly, ratings are along dimensions of pleasantness, attractiveness, sexiness, intensity or masculinity-femininity (e.g. Allen, Cobey, Havlíček, & Roberts, 2016; Gildersleeve, Haselton, Larson, & Pillsworth, 2012). For example, in a study investigating changes in body odour across the menstrual cycle, it was found that men rated women's odour as most sexually attractive when they were mid-cycle, when conception probability peaks (Kuukasjärvi et al., 2004). This is an important and interesting finding, and the term 'sexually attractive' is clearly useful and practical in that it allows us to investigate changes in mating-relevant qualities, however, it gives us no specific information regarding the changes in the perceptual quality of these body odours; in other words, it does not tell us what sexually attractive odours smell like. Additionally, while research has found there to be sex differences in both volatile axillary compounds (Penn et al., 2007) and the ratios of certain non-volatile compounds (Troccaz et al., 2009), these do not always appear to be reflected in perceptual ratings of masculinity and femininity of odours. For instance, Mutic and colleagues (2015) found that odours were rated as mostly masculine, regardless of the donors' actual sex, suggesting that these terms may not adequately capture the relevant perceptual differences between odours.

How then can we improve upon the ratings of the perceptual qualities of odours and increase the ecological validity of our measures? One solution would be to utilise a 'bottom-up' approach to identify dimensions along which people tend to categorise odours which can then be combined into a new lexicon for odour description. With this aim in mind, it may be beneficial to develop and utilise such a lexicon with those who have experience and training in odour evaluation – namely perfumers and fragrance evaluators. Perhaps they can provide us with more detailed descriptions of odours, allowing us to further investigate the potentially fine-grained differences between individual odours, and thus their role in human social interaction.

Research following this line of investigation, while uncommon, does show some promise. One study found that, while there was no difference in hedonic ratings of odours given by laymen and trained perfumers, perfumers gave richer verbal descriptions of odours (Sezille, Fournel, Rouby, Rinck, & Bensafi, 2014). Additionally, Wedekind and colleagues (2007) found that trained perfumers were capable of describing human body odours in such a way that highly variable genetic information (major histocompatibility allelic specificity) could be distinguished, but untrained assessors could not. More recently, Troccaz and colleagues (2015) trained assessors in verbally describing certain chemical compounds which appear in human axillary odours. Their main aim was to elucidate the perceptual and microbiotic variation between individuals who use or do not use antiperspirants, but the findings also revealed some sex differences in the perceptual qualities of non-treated odours. Male odours tended to receive higher ratings of acid-spicy odour intensity than female odours, although this was only statistically significant in some men. These findings suggest then that olfactory training and experience with assessing odours, such as that gained by perfumers, may lead to more accurate descriptions of odours than can be achieved by non-trained assessors.

The aim of the current study was therefore to explore the different dimensions of body odours which are perceived and to utilise these to establish a lexicon which could be used to describe some qualitative components of body odours, beyond simple hedonic descriptors. A panel of perfumers and fragrance evaluators worked together on an exemplar set of axillary odours to compile a list of verbal descriptors for qualities of these odours. They then assessed the presence and intensity of each of these qualities in a set of axillary (armpit) odours from male and female odour donors. To test the utility of these assessments and this lexicon in discriminating known differences between these individuals, we evaluated whether scores on these descriptors reliably predicted the sex of odour donors, since we know that sex can be identified based on the chemical compounds present in axillary odours (Penn et al., 2007; Schleidt, 1980; Troccaz et al., 2009).

Materials and Methods

The study was approved by the University of Stirling ethical review board and all donors gave written consent before taking part in the study.

Odour Donors

We recruited heterosexual individuals only as previous studies have found that odour quality differs with sexual orientation (Martins, Preti, Crabtree, Runyan, Vainius, & Wysocki, 2005). In total sixty-two individuals (31 women) were recruited to provide odour samples (mean age of women = 28, SD = 8.59, range 20-51 years; mean age of men = 29.47, SD = 9.21, range 20-51 years). In line with previous research (Roberts, Havlíček, & Petrie, 2013), we instructed our donors to avoid drinking alcohol, being in smoky places, exercising and eating certain strong-smelling foods (e.g. garlic, asparagus, curry) one day prior to, and during, odour collection periods. They were additionally asked to refrain from sexual activity and to avoid sharing their bed with anyone during the odour collection phases (Kohoutová, Rubešová, & Havlíček, 2011; Lenochová et al., 2012; Roberts et al., 2011). Donors were provided with fragrance free soap (Simple Pure™) and asked to use only this in place of any fragranced hygiene products for 24 hours prior to odour collection.

Each individual underwent one 24 hour odour collection period. Each donor was provided with 100% cotton oval shaped make-up pads (approximately 9.5cm x 6.5cm, 3mm thick, Cosmetic Oval Pads, The Boots Company PLC) and surgical tape (Finepore™, 2.5cm wide). Donors were instructed to apply the cotton pad onto their armpit, using the tape to hold this in place, and to remove it 24 hours later. The donors returned the samples, labelled and in sealed plastic bags, to the lab within 2 hours of removal, where they were stored in a freezer at -30˚C until use. Samples were thawed at room temperature for 2 hours prior to test sessions and re-frozen between test sessions. Previous research suggests freezing and thawing of samples has minimal impact on the perceptual quality of the odour (Lenochová, Roberts, & Havlíček, 2009; Roberts, Gosling, Carter, & Petrie, 2008).

Odour Assesors

Two perfumers (1 male and 1 female) and two fragrance evaluators (both female) volunteered to take part in the study. They were aged 29-45 (mean = 38.25, SD = 7.27) and had been working in the industry for between 6-18 years (mean = 11.75, SD = 5.05). Perfumers and fragrance evaluators typically work together to meet client briefs for fragrances. Evaluators are heavily involved in smelling the fragrances, in order to ascertain if these meet the brief, but it is the perfumer who is responsible for designing the fragrance, and as such perfumers have more knowledge of raw ingredients and more years of training.

Procedure

As a group, the assessors evaluated ten axillary samples (from five men and five women, of the original 62 donors recruited) and together drew up a list of 15 basic descriptors present in these samples. Descriptors were taken from standard “olfactive maps” used throughout the fragrance industry to describe and map odors for commercial investigations. Fragrance houses create their own maps and use these internally to train and calibrate their experts. As our experts were all part of the same team they were easily able to agree on definitions of descriptors. The descriptors chosen were known to all the experts and have definitions in olfactory terms (see Table 1). These were Musty, Mouldy, Earthy, Onion, Spicy, Fatty, Oily, Greasy, ChipFat, Animalic, Vegetable, Heavy, Milky, Sweet, and Metallic. Having established and agreed upon this common semantic inventory, they then smelled each of the 62 samples (including the 4 which had been used for the initial evaluation, and blind to the donors’ identity and sex) and rated each sample according to each descriptor using a 10-point scale of intensity (0 = no presence of this descriptor, 10 = extreme presence of descriptor). The category ‘other’ was also included to allow for the possibility that important descriptors may have been missed from the original list. The category ‘other’ was only used 11 times across all samples and assessors (out of a possible 248 ratings). No single descriptor came out of the ‘other’ category; ‘other’ descriptors used were: Green (1), Chocolate (3), Salty (1), Cumin (1), Grass (1), Maltol (1), Cheese (1), Cotton (1) and Sharp (1). The low frequency of use of this category, and the lack of a common new descriptor emerging from the larger set of samples, suggests that the original 15 descriptors were robust and comprehensive. Additionally, for each odour sample, the assessors provided an explicit judgment of whether they thought it was from a man or a woman.

Each of the four assessors smelled all 62 of the samples over the space of two weeks. Samples were rated in groups of 5 (and one group of 2), with assessors rating no more than 10 samples in a day. Sets of samples were removed from the freezer and allowed to defrost before use, then removed from the bags and assessed straight from the cotton pad. All four assessors completed their assessments of each set during the same day. Ratings were completed in the same room at the perfumers' place of work.

Results

Exploratory factor analysis of Lexicon

To control for differences in the use of the scale across assessors, each assessor's individual scores for each descriptor were standardised by computing *z*-scores. It should be noted that each assessor had one descriptor which they never detected within any of the samples – one assessor never detected any *Mouldy* odours, another assessor never detected any *Animalic* odours, and the final two assessors never detected any *Metallic* odours. Intraclass correlation coefficients (ICC) are a standard method for assessing reliability and agreement of ratings (Shrout & Fleiss, 1979) and were conducted in order to establish the inter-assessor reliability across the scale. As can be seen from Table 2, six of the fifteen descriptors had ICC's above .4 (.40-.59 = fair, .60-.74 = good, > .74 = excellent, Cicchetti & Sparrow, 1981; Fleiss, 1981). These were *Onion*, *Spicy*, *Animalic*, *Heavy*, *Milky* and *Sweet*. To explore the underlying structure of our lexicon and the semantic dimensions within this, we conducted a factor analysis using only the 6 descriptors that showed good inter-rater reliability as measured via intraclass correlation coefficients (Table 2). Suitability of the 6 items for factor analysis was initially examined, using several well recognised criteria.

First, all 6 items were found to be somewhat correlated ($r > .3$) with at least one other item (Table 3). Second, the Kaiser-Meyer-Olkin measure of sampling adequacy (.806) was above the recommended value of .6, and Bartlett's test of sphericity was significant, $x^2(15) = 148.46$, $p < .001$. Furthermore, the diagonals of the anti-image correlation matrix were all found to be over .5, and finally all variables had communalities above .3, suggesting common variance with other items. These analyses suggest the data are suited to factor analysis.

We calculated mean z-scores for each of the 6 descriptors and for each donor, and then conducted an exploratory factor analysis (principal axis factoring) using varimax rotation. After rotation eigenvalues showed that the total variance explained by factors one and two was 40.42% and 20.19% respectively, with this two factor solution explaining 60.62% of the total variance. All 6 items had primary factor loadings above .4, and only one was found to cross-load onto another factor at above .3 (*Onion*), but this was deemed acceptable as the primary factor loading was high (.753), so all 6 variables were retained and two factors were extracted from the model; Spicy/Animalic and Sweet/Milky (Table 4).

Identifying sex from odour

Binomial tests were used to compare the observed frequency of correct explicit judgments (assessors' guesses of odour donor's sex; Figure 1) against that expected by chance (.5). Only assessor 1 was capable of correctly inferring the sex of the samples at a significantly above chance level, p = .003 (69% correct), with assessor 3 showing only a marginal significance, p = .056 (63% correct) and assessors 2 and 4 performing at a close to chance level: assessor 2, p = .374 (56% correct); assessor 4, p = .899 (52% correct).

Ratings and sex of odour

We then investigated differences in descriptor ratings between male and female odours. We calculated the mean z-score from all assessors for each donor, for each descriptor. A repeated measures ANOVA was conducted, with descriptor as the within-subjects factor (15 levels) and donor sex as the between-subjects factor. There was no main effect of descriptor, F (14, 840) < .01, reflecting the fact we use standardised scores to control for potential differences in raters' use of the rating scale, but there was a significant interaction between descriptor ratings and donor sex, F (14, 840) = 1.789, p = .036. Post hoc independent samples t-tests revealed that there were significant differences between male and female odours in rating of

Spicy, *Animalic* and *Metallic*, with men receiving higher ratings for all three of these descriptors (Table 5), though it must be noted that only *Spicy* and *Animalic* received acceptably high intraclass correlation coefficients (Table 2).

Following on from this we computed composite scores for each donor for each of the two extracted factors (Spicy/Animalic and Sweet/Milky) and independent samples *t*-tests were conducted to compare factor scores between male and female odours. There was no significant difference between male and female odours on Sweet/Milky scores (factor 2), t (60) = .36, p = .724, but there was a significant difference in scores on Spicy/Animalic (factor 1), t (60) = 2.23, p = .029, with men scoring higher in this factor than women (Figure 2).

Discussion

Hedonic evaluation of individual variation in body odours detected by humans is almost always limited to assessment on a small number of scales, many of which do not focus on specific qualities of the odour percept. While these scales do provide useful measures, they inevitably miss much of the diversity and complexity in human body odours, which contain hundreds of unique volatile compounds in individually variable patterns of abundance. The main aim of this study was to explore the development of a more detailed set of body odour descriptors which better capture this diversity, with the aim of creating a new lexicon for body odour description. We initially used 15 descriptors, although only 6 were used consistently across our trained assessors. This perhaps reflects the difficulty in describing odour even for trained professionals, but nonetheless suggests that these 6 descriptors may be capturing important odour qualities. To validate the utility of these descriptors, we tested whether they differentiated between donor sex, finding that scores on the descriptors *Spicy*, *Animalic* and *Metallic* were each significantly higher in male samples than in female samples. We also used factor analysis to further explore the odour evaluations, which revealed a two factor structure to the data. We found that Spicy/Animalic scores were significantly higher in male than female odours. Our findings indicate that this novel lexicon is a useful tool for the description of human body odour variation.

We found that male odours received significantly higher ratings of three descriptors in our study. The result for the descriptor *Spicy* is consistent with the sex differences in Spicy ratings found by Troccaz and colleagues (2015), and the significant sex differences in *Animalic* and *Metallic* descriptor scores further extends this. Our exploratory factor analysis generated two factors, the first (Spicy/Animalic) comprising the descriptors *Onion*, *Spicy*, *Animalic* and *Heavy*, and the second (Sweet/Milky) containing the descriptors *Milky* and *Sweet*. Our analyses revealed a significant difference between men and women's Spicy/Animalic scores, in keeping with the single-descriptor differences for *Spicy* and *Animalic* (higher scores in male odours), and incorporating also the descriptors *Onion* and *Heavy*, both of which scored more highly in male odours (though not significantly so) in the single descriptor ratings.

Given the finding above, that there appear to be perceptual differences in male and female odours (Doty, Orndorff, Leyden, & Kligman, 1978; Hold & Schleidt, 2010; Russell, 1976; Schleidt, 1980), and other findings showing that there are chemical differences between male and female body odours (Penn et al., 2007; Troccaz et al., 2009), we were surprised that our assessors were not all successful at discriminating sex of the odour donors at above chance levels. Only one assessor appeared to be able to do this reliably, with another's success rate being almost better than chance, and two performing at chance levels. However, to date, the literature on sex discrimination of axillary odours is ambiguous, with reported success rates varying considerably, ranging from 20%-100% of participants (Doty, Orndorff, Leyden, & Kligman, 1978; Hold & Schleidt, 2010; Russell, 1976; Schleidt, 1980). We believed that the fragrance expertise our olfactory assessors had would benefit their performance on this task, though that was not the case, and coupled with this variance in performance noted in the literature, suggests that conscious sex categorisation of axillary odours is not a straightforward task.

Our lexicon was successful at quantifying sex differences in axillary odours, despite mixed success in sex identification in the assessors' explicit judgments. Future research should now focus on investigating the evaluation of other traits, both stable and those which fluctuate, that appear to be cued in body odour. These may be related to other single descriptors, or different combinations of descriptors, or even relating to the factors extracted from our exploratory analysis. For example, although the Sweet/Milky scores from our factor analysis did not distinguish between male and female odours, the contributing descriptors (*Milky* and *Sweet*) might be correlated with some other important social attribute, such as personality characteristics or fertility.

The verbal classification of odours is inherently difficult. Often expressions relating to the source of an odour from another modality (e.g., taste – sweet) are employed to tackle this (Kaeppler & Mueller, 2013). These individual odour classification systems based on perceptual characteristics vary greatly across studies and do not tend to converge into one generally accepted system. Nevertheless, numerous specifically designed classification systems have been developed, often for practical reasons, for example for sensory assessment of food products such as wine (Noble et al., 1984), coffee (Williams & Arnold, 1985) or cosmetic products such as perfumes. For instance, perfumers commonly use the OSMOZ system (see http://www.osmoz.com/encyclopedia/olfactory-groups), which classifies fragrances into 10 main categories, each of which further consists of four subcategories. Such a system allows for the relatively easy classification of odours which captures relatively fine nuances between individual fragrances and has been successfully used in research on perfume selection (Sobotková, Fialová, Roberts, & Havlíček, 2016). Here we aimed to develop a similar tool specifically tailored for body odours. To do so, we employed a bottom-up approach while utilising descriptors used by professional perfumers who are expected to have a richer odour-related vocabulary. An alternative approach was recently employed by Troccaz et al. (2009) who trained their evaluators in identification of chemical compounds characteristic of body odour. The main limitation of this approach is that the body odour may have different perceptual qualities as compared to its components. This is primarily a consequence of the emerging perceptual qualities which arise from the complex nature of body odours (Laing, 1994). However, there is a potential disadvantage to our approach, such that we had only a small number of assessors who may not have fully captured the whole range of suitable body odours descriptors. In order to minimise the impact of this we allowed them to use further descriptors while they were rating the full set of the body odour samples, and in support of our lexicon we found that additional descriptors were only rarely, and not consistently, used. It should also be noted that only six out of our fifteen original descriptors showed acceptable internal consistency. This may be a result of the small number of olfactory experts used in this study, due to the limited access to these individuals, but it could also indicate that even among professionals there is a high level of idiosyncrasy in odour perception. Nevertheless, future studies should aim to build on and extend this work by employing a broader set of assessors and including more thorough calibration and practice sessions to truly investigate the utility of our lexicon. It would also be valuable to test the lexicon with lay individuals as such research could also potentially allow participants to use their own descriptors which may capture some unique descriptors missed in the current study. Future research may also benefit from investigating whether there are sex differences in the use of our lexicon as there is evidence of sex differences in olfactory performance (Brand & Millot, 2001) which may affect this.

The lexicon developed here will not only be of benefit to researchers, but also potentially for the fragrance industry. Our approach could be useful for categorising body odours for practical purposes, for example, as a way to classify individual body odours in order to explore how certain fragrance ingredients or fragrance accords interact with and complement different body odour categories. It is known that some individuals choose fragrances that complement their own body odour, while others aren't as good at choosing fragrances; the same fragrance mixed with a different body odour can produce an odour blend that smells worse than the body odour by itself (Lenochova et al., 2012). Additionally, it was recently found that individually selected fragrances promote individual discrimination compared to allocated fragrances (Allen, Havlíček, & Roberts, 2015). Choosing the "right" fragrance is clearly difficult for some people, and categorising body odour and investigating which fragrances complement given odour categories could offer a potential practical solution in the development of tailored perfumes

We also suggest that psychological research into human olfactory communication could benefit greatly from this kind of nuanced measure of the perceptual qualities of odours, beyond the limited set of rating scales (e.g. pleasantness, attractiveness, intensity) used to date. In this regard, the main challenge ahead is now to establish whether this lexicon can also be successfully used by non-perfumers, given that it was developed by individuals with unusual levels of olfactory expertise. It seems likely that some of the descriptors used here will be familiar to untrained individuals (e.g. sweet, spicy, heavy), and so perhaps with training and further standardisation of descriptor definitions there may be scope to incorporate these descriptors into future research working with lay individuals.

In conclusion, our study presents the first attempt to explore dimensions along which human body odours can be classified. A similar approach has been previously used for facial perception, finding that the main dimensions include sex, attractiveness, trustworthiness, dominance and age (for details see Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Our study indicates that the dimensions employed for body odour classification considerably differ from facial perception. However, generalisability of our findings across different social contexts and populations remains to be explored by future studies. The novel lexicon presented here is potentially a useful tool for improving our ability to measure the perceptual quality of body odours. Future research is needed to work on integrating molecular chemistry and human olfactory perception in order to fully appreciate the range and variation within human body odours, and the role that these may serve in human social interactions.

Conflict of interest

KW was employed by Seven Scent Ltd. and was President of the British Society of Perfumers when the study was conducted. However, neither role introduces any conflict of interest with the specific nature of this study.

References

Allen, C., Cobey, K. D., Havlíček, J., & Roberts, S. C. (2016). The impact of artificial fragrances on the assessment of mate quality cues in body odor. *Evolution and Human Behavior, 37*(6), 481–489.

Allen, C., Havlíček, J., & Roberts, S. C. (2015). Effect of fragrance use on discrimination of individual body odor. *Frontiers in Psychology*, *6*, 1115.

Brand, G., & Millot, J. L. (2001). Sex differences in human olfaction: Between evidence and enigma. *The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*, *54*, 259–270.

Chen, D., & Haviland-Jones, J. (2000). Human olfactory communication of emotion. *Perceptual and Motor Skills*, *91*, 771–781.

Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing the interrater reliability of specific items in a given inventory. *American Journal of Mental Deficiency*, *86*, 127–137.

Doty, R. L., Orndorff, M. M., Leyden, J. J., & Kligman, A. (1978). Communication of gender from human axillary odors: relationship to perceived intensity and hedonicity. *Behavioral Biology*, *23*, 373–80.

Fialová, J., Roberts, S. C., & Havlíček, J. (2016). Consumption of garlic positively affects hedonic perception of axillary body odour. *Appetite*, *97*, 8–15.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

Gildersleeve, K. A., Haselton, M. G., Larson, C. M., & Pillsworth, E. G. (2012). Body odor attractiveness as a cue of impending ovulation in women: Evidence from a study using hormone-confirmed ovulation. *Hormones and Behavior*, *61*, 157–166.

Havlíček, J., Dvorakova, R., Bartos, L., & Flegr, J. (2006). Non-advertized does not mean concealed: Body odour changes across the human menstrual cycle. *Ethology*, *112*, 81–90.

Havlicek, J., & Lenochova, P. (2006). The Effect of Meat Consumption on Body Odor Attractiveness. *Chemical Senses*, *31*(8), 747–752.

Havlíček, J., & Roberts, S. C. (2009). MHC-correlated mate choice in humans: A review. *Psychoneuroendocrinology*, *34*, 497–512.

Hold, B., & Schleidt, M. (2010). The Importance of Human Odour in Non-verbal Communication. *Zeitschrift Für Tierpsychologie*, *43*(3), 225–238.

Kaeppler, K., & Mueller, F. (2013). Odor Classification: A Review of Factors Influencing Perception-Based Odor Arrangements. *Chemical Senses*, *38*(3), 189–209.

Kohoutová, D., Rubešová, A., & Havlíček, J. (2011). Shaving of axillary hair has only a transient effect on perceived body odor pleasantness. *Behavioral Ecology and Sociobiology*, *66*, 569–581.

Kuukasjärvi, S., Eriksson, C. J. P., Koskela, E., Mappers, T., Nissinen, K., & Rantala, M. J. (2004). Attractiveness of women's body odors over the menstrual cycle: the role of oral contraceptives and receiver sex. *Behavioral Ecology*, *15*, 579–584.

Laing, D. G. (1994). Perceptual odour interactions and objective mixture analysis. *Food Quality and Preference*, *5*(1–2), 75–80.

Lenochová, P., Vohnoutová, P., Roberts, S. C., Oberzaucher, E., Grammer, K., & Havlíček, J. (2012). Psychology of fragrance use: perception of individual odor and perfume blends reveals a mechanism for idiosyncratic effects on fragrance choice. *PLoS One*, *7*, e33810.

Martins, Y., Preti, G., Crabtree, C. R., Runyan, T., Vainius, A. A., & Wysocki, C. J. (2005). Preference for human body odors is influenced by gender and sexual orientation. *Psychological Science*, *16*, 694–701.

Moshkin, M., Litvinova, N., Litvinova, E. A., Bedareva, A., Lutsyuk, A., & Gerlinskaya, L. (2012). Scent recognition of infected status in humans. *The Journal of Sexual Medicine*, *9*, 3211–3218.

Mutic, S., Moellers, E. M., Wiesmann, M., & Freiherr, J. (2015). Chemosensory Communication of Gender Information: Masculinity Bias in Body Odor Perception and Femininity Bias Introduced by Chemosignals During Social Perception. *Frontiers in Psychology*, *6*, 1980.

Noble, A. C., Arnold, R. A., Masuda, B. M., Pecore, S. D., Schmidt, J. O., & Stern, P. M. (1984). Progress Towards a Standardized System of Wine Aroma Terminology. *American Journal of Enology and Viticulture*, *35*, 107–109.

Penn, D. J., Oberzaucher, E., Grammer, K., Fischer, G., Soini, H. A., Wiesler, D., … Brereton, R. G. (2007). Individual and gender fingerprints in human body odour. *Journal of the Royal Society Interface*, *4*, 331–340.

Rikowski, A., & Grammer, K. (1999). Human body odour, symmetry and attractiveness. *Proceedings of the Royal Society B: Biological Sciences*, *266*, 869–874.

Roberts, S. C., Gosling, L. M., Spector, T. D., Miller, P., Penn, D. J., & Petrie, M. (2005). Body odor similarity in noncohabiting twins. *Chemical Senses*, *30*, 651–656.

Roberts, S. C., Havlíček, J., & Petrie, M. (2013). Repeatability of odour preferences across time. *Flavour and Fragrance Journal*, *28*, 245–250.

Roberts, S. C., Kralevich, A., Ferdenzi, C., Saxton, T. K., Jones, B. C., DeBruine, L. M., … Havlíček, J. (2011). Body odor quality predicts behavioral attractiveness in humans. *Archives of Sexual Behavior*, *40*, 1111–1117.

Russell, M. J. (1976). Human olfactory communication. *Nature*, *260*, 520–522.

Schleidt, M. (1980). Personal odor and nonverbal communication. *Ethology and Sociobiology*, *1*, 225–231.

Sezille, C., Fournel, A., Rouby, C., Rinck, F., & Bensafi, M. (2014). Hedonic appreciation and verbal description of pleasant and unpleasant odors in untrained, trainee cooks, flavorists, and perfumers. *Frontiers in Psychology*, *5*, 12.

Sobotková, M., Fialová, J., Roberts, S. C., & Havlíček, J. (2016). Effect of Biological Relatedness on Perfume Selection for Others: Preliminary Evidence. *Perception*, 30100661668251.

Sorokowska, A., Sorokowski, P., & Szmajke, A. (2012). Does Personality Smell? Accuracy of Personality Assessments Based on Body Odour. *European Journal of Personality*, *26*, 496–503.

Shrout, P.E. & Fleiss, J.L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86* (2), 420-428.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, *66*(1), 519–545.

Troccaz, M., Borchard, G., Vuilleumier, C., Raviot-Derrien, S., Niclass, Y., Beccucci, S., & Starkenmann, C. (2009). Gender-specific differences between the concentrations of nonvolatile (R)/(S)-3-methyl-3-sulfanylhexan-1-Ol and (R)/(S)-3-hydroxy-3-methyl-hexanoic acid odor precursors in axillary secretions. *Chemical Senses*, *34*(3), 203–10.

Troccaz, M., Gaïa, N., Beccucci, S., Schrenzel, J., Cayeux, I., Starkenmann, C., & Lazarevic, V. (2015). Mapping axillary microbiota responsible for body odours using a culture-independent approach. *Microbiome*, *3*(1), 3.

Wedekind, C., Escher, S., Van de Waal, M., & Frei, E. (2007). The major histocompatibility complex and perfumers ’ descriptions of human body odors. *Evolutionary Psychology*, *5*, 330–343.

Wedekind, C., Seebeck, T., Bettens, F., & Paepke, A. J. (1995). MHC - dependent mate preferences in humans. *Proceedings: Biological Sciences*, *260*, 245–249.

Williams, A. A., & Arnold, G. M. (1985). A comparison of the aromas of six coffees characterised by conventional profiling, free-choice profiling and similarity scaling methods. *Journal of the Science of Food and Agriculture*, *36*(3), 204–214.

Winternitz, J., Abbate, J. L., Huchard, E., Havlíček, J., & Garamszegi, L. Z. (2017). Patterns of MHC-dependent mate selection in humans and nonhuman primates: a meta-analysis.

Molecular Ecology.

Table 1. Definitions of the 15 descriptors used by evaluators and perfumers in body odor assessment

Descriptor	Agreed definition
Musty	Stale air, old furniture
Moldy	Household mold, mold found on clothes, bread mold (not cheese mold)
Earthy	Soil, wet forest floor, mud, wet tree bark
Onion	The smell of raw onion, red, white, spring and leeks
Spicy	Refers only to culinary spices such as clove, nutmeg, cumin, anise, pepper, etc.
Fatty	Cold fats and oils used for cooking including butter and lard, margarine, olive oil, vegetable oil, and rendered beef fat
Oily	Oil paint, violet leaf absolute, car engine oil, WD40, non-edible oils
Greasy	Dirty human scalp and/or hair
Chipfat	Fat from a deep fat fryer used to cook potato
Animalic	Odors from an animal source including goat, horse, sweat, skin, fur, leather, etc.
Vegetable	Savory vegetable aroma, vegetable stock or soup, cooked vegetables, raw vegetables including potato, carrot, celery
Heavy	Non-volatile odors, similar olfactive feel to larger musk molecules
Milky	Lactonic, milk from all animal sources
Sweet	Vanilla, chocolate, sugar
Metallic	Smells like metal, hot metal, tin, iron

Table 2 Intraclass correlation coefficients (ICC) for the 4 assessors' *z*-score ratings across the 15 descriptors (not including 'other'). 95% confidence intervals are shown. ICI values above .4 are deemed acceptable and are indicated in bold.

Descriptor	ICC Z scores	95% CI lower bound	95% CI upper bound
Musty	.155	-.249	.453
Mouldy	-.043	-.590	.338
Earthy	.080	-.361	.404
Onion	**.552**	**.338**	**.710**
Spicy	**.589**	**.393**	**.734**
Fatty	-.135	-.679	.265
Oily	.160	-.242	.456
Greasy	.301	-.034	.547
Chipfat	.324	.001	.562
Animalic	**.531**	**.284**	**.702**
Vegetable	-.281	-.894	.171
Heavy	**.598**	**.405**	**.740**
Milky	**.475**	**.224**	**.660**
Sweet	**.633**	**.457**	**.762**
Metallic	-.155	-.917	.304

Table 3 Correlations between the 6 descriptors which were included in the factor analysis.

	Onion	Spicy	Animalic	Heavy	Milky
Spicy	.703				
Animalic	.549	.568			
Heavy	.635	.700	.546		
Milky	-.268	-.285	-.171	-.105	
Sweet	-.461	-.386	-.313	-.255	.522

Table 4 Loadings and communalities for the 6 descriptor items based on mean *z*-scores from the 4 assessors.

Descriptor	**Factor 1 (Spicy/Animalic)**	**Factor 2 (Sweet/Milky)**	**Communalities**
Onion	**.753**	-.328	.675
Spicy	**.815**	-.265	.735
Animalic	**.645**	-.180	.448
Heavy	**.836**	-.042	.701
Milky	-.095	**.665**	.451
Sweet	-.263	**.747**	.627

Table 5 Mean standardised scores for each descriptor for male and female samples. *p* values are taken from post hoc independent samples *t*-tests. Significant values are shown in bold.

Descriptor	**Male mean rating**	**Female mean rating**	***p***
Musty	.0094	-.0094	.891
Mouldy	.0616	-.0616	.260
Earthy	-.0175	.0175	.792
Onion	.0670	-.0670	.424
Spicy	**.1782**	**-.1782**	**.035**
Fatty	.0150	-.0150	.806
Oily	-.0879	.0879	.197
Greasy	-.0936	.0936	.197
ChipFat	-.0502	[illegible]502	.497
Animalic	**.1919**	**-.1919**	**.004**
Vegetable	-.0940	.0940	.104
Heavy	.1471	-.1471	.085
Milky	.0039	-.0039	.961
Sweet	.0058	-.0058	.948
Metallic	**.0689**	**-.0689**	**.044**

Figure 1. Proportion of correct explicit judgments of donor sex by each assessor. Assessors 1 and 4 evaluators. Assessors 2 and 3 are perfumers. Dashed line indicates chance level. ** $p < .01$

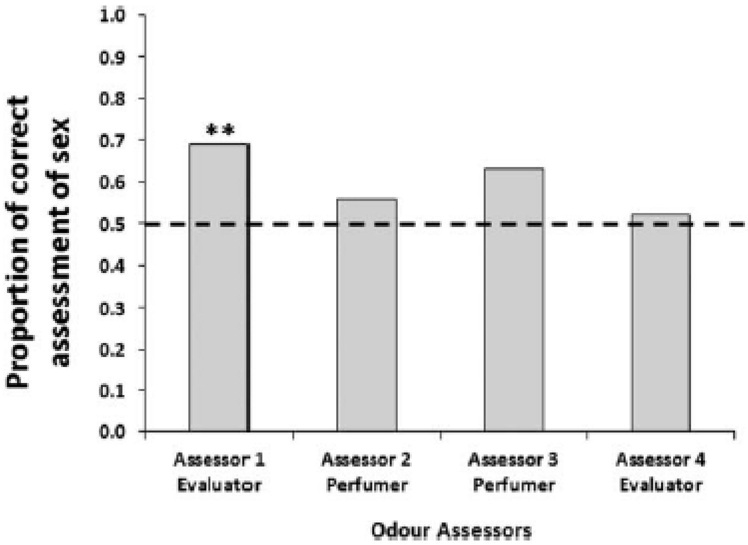

Figure 2. Mean ratings for males and females for the factors generated from the factor analysis. Error bars represent ± 1 SEM. * $p < .05$